

明 細 書

オーディオ信号符号化装置、オーディオ信号復号化装置、方法、及びプログラム

技術分野

[0001] 本発明は、オーディオ信号の符号化装置、復号化装置、方法、および、プログラムに関するものである。

背景技術

[0002] 従来のオーディオ信号符号化方法、および、復号化方法としては、公知なものとしてISO/IECの国際標準方式、通称MPEG方式などが挙げられる。現在、幅広い応用を持ち、低ビットレート時でも高音質な符号化方式として、ISO/IEC 13818-7、通称MPEG2 AAC (Advanced Audio Coding)などがあげられる。本方式の拡張規格も複数規格化が現在なされている。

[0003] その一つとして、空間音響情報(Spatial Cue Information)もしくは、聴覚的音響情報(Binaural Cue)と呼ばれる情報を利用する技術がある。このような技術の例としては、ISO国際標準規格であるMPEG-4 Audio (ISO/IEC 14496-3)において定められたパラメトリックステレオ(Parametric Stereo)方式がある。また、別の例として、米国公開特許US2003/0035553 “Backwards-compatible Perceptual Coding of Spatial Cues”において開示される方式がある。(非特許文献1参照)。また、別の方式も提案されている(例えば、特許文献1、特許文献2等参照)。

非特許文献1: ISO/IEC 14496-3:2001 AMD2 “Parametric Coding for High Quality Audio”

特許文献1: 米国公開特許US2003/0035553 “Backwards-compatible Perceptual Coding of Spatial Cues”

特許文献2: 米国公開特許US2003/0219130 “Coherence-based Audio Coding and Synthesis”

発明の開示

発明が解決しようとする課題

[0004] しかしながら、従来のオーディオ信号符号化方法、および、復号化方法では、例えば背景技術に記載のAACなどでは、マルチチャンネルの信号を符号化する際に、チャンネル間の相関を十分に生かしきれていないため、低ビットレート化することが困難であった。チャンネル間の相関を用いて符号化を実施する場合においても、人間の音源の方向知覚の特性や、拡がり感に関する特徴をもちいることで得られる符号化効率の向上などの効果を、十分に量子化と符号化に活かしかれていないという課題があった。

[0005] また従来の方法では、マルチチャンネルの信号を符号化したものを復号化する際に、2つのスピーカやヘッドホンなどで再生する場合において、一度、すべてのチャンネルを復号化し、その後、ダウンミックスなどの方法を用いて、前記2つのスピーカやヘッドホンで再生すべきオーディオ信号を加算により生成しなければならなかった。このことは2つのスピーカやヘッドホンで再生する場合に、多くの計算量や計算用のバッファを要し、ひいては、それを実装するDSPなどの計算手段の消費電力やコストを高める原因となった。

[0006] 本発明は、上記課題を解決するものであって、マルチチャンネルの信号を符号化する際の符号化効率を向上するオーディオ信号符号化装置、その装置から得られた符号を復号化するオーディオ信号復号化装置を提供することを目的とする。

課題を解決するための手段

[0007] 本発明のオーディオ信号符号化装置は、複数のチャンネルそれぞれの原音信号を、前記原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とに符号化するオーディオ信号符号化装置であって、前記原音信号を混入して得た混入信号を符号化することによって前記混入信号情報を生成する混入信号符号化手段と、前記原音信号に基づいて前記特徴量を算出すると共に、各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる符号化方法を決定し、前記決定された符号化方法を用いて前記算出された特徴量を符号化することによって前記補助情報を生成する補助情報生成手段とを備える。

- [0008] また、前記補助情報生成手段は、異なる量子化精度が得られる量子化点を定義する複数のテーブルを予め記憶しており、前記特徴量を、前記複数のテーブルのうちの前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つによって定義される量子化点に量子化することによって符号化してもよい。
- [0009] また、前記補助情報生成手段は、前記原音信号間のレベル差及び位相差の少なくとも一方を算出して前記特徴量としてもよく、さらに、前記算出されたレベル差及び位相差に基づいて、リスナーが知覚すると推定される音像の方向を算出して前記特徴量としてもよい。
- [0010] また、前記補助情報生成手段は、リスナーの正面方向から左右対称に設けられる量子化点を定義する第1のテーブルと、リスナーの左方向から前後非対称に設けられる量子化点を定義する第2のテーブルとを予め記憶しており、前記チャンネル情報がリスナーの左前方と右前方とを示す場合には、前記特徴量を前記第1のテーブルによって定義される量子化点に量子化することによって符号化し、前記チャンネル情報がリスナーの左前方と左後方とを示す場合には、前記特徴量を前記第2のテーブルによって定義される量子化点に量子化することによって符号化してもよい。
- [0011] また、前記補助情報生成手段は、前記原音信号間の類似度を算出して前記特徴量としてもよく、また、前記原音信号間の相互相関値、又はその絶対値を、前記類似度として算出してよく、さらに、前記算出された類似度に基づいて、リスナーが知覚すると推定される音像の拡がり感及び距離感の少なくとも一方を算出して前記特徴量としてもよい。
- [0012] 上記課題を解決するため、本発明のオーディオ信号復号化装置は、複数のチャンネルそれぞれの原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とを、前記複数のチャンネルそれぞれの再生信号に復号化するオーディオ信号復号化装置であって、各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる復号化方法を決する復号化方式切り替え手段と、前記決定された復号化方法を用いて前記補助情報を前記特徴量に復号化する信号間情報復号手段と、前記混入信号情報と前記

復号化された特徴量とを用いて、前記複数のチャンネルそれぞれの再生信号を生成する信号合成手段とを備える。

- [0013] また、前記補助情報は、前記特徴量を、異なる量子化精度が得られる量子化点を定義する複数のテーブルのうち前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つによって定義される量子化点に量子化することによって符号化されており、前記信号間情報復号手段は、前記複数のテーブルを予め記憶しており、前記複数のテーブルのうちの前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つを用いて、前記補助情報を前記特徴量に復号化してもよい。
- [0014] また、前記特徴量は、前記原音信号間のレベル差、位相差、及び、リスナーが知覚すると推定される音像の方向の少なくとも一つを表し、前記信号間情報復号手段は、リスナーの正面方向から左右対称に設けられる量子化点を定義する第1のテーブルと、リスナーの左方向から前後非対称に設けられる量子化点を定義する第2のテーブルとを予め記憶しており、前記チャンネル情報がリスナーの左前方と右前方とを示す場合には、前記第1のテーブルを用いて前記補助情報を前記特徴量に復号化し、前記チャンネル情報がリスナーの左前方と左後方とを示す場合には、前記第2のテーブルを用いて前記補助情報を前記特徴量に復号化してもよい。
- [0015] また、前記特徴量は、前記原音信号間のレベル差、位相差、類似性、及び、リスナーが知覚すると推定される音像の方向、拡がり感、並びに距離感の少なくとも一つを表すとしてもよい。
- [0016] また、前記信号合成手段は、前記特徴量が、前記原音信号間のレベル差、位相差、及び類似性の少なくとも一つを表す場合に、前記混入信号情報によって表される音声信号に前記特徴量に応じたレベル差、位相差、及び類似性を与えることによって、前記再生信号を生成してもよい。
- [0017] また、本発明は、このようなオーディオ信号符号化装置及びオーディオ信号復号化装置として実現できるだけでなく、このような装置が備える特徴的な手段によって実行される処理をステップとする方法として実現することも、また、それらのステップをコンピュータに実行させるプログラムとして実現することもできる。そして、そのようなプログ

ラムは、CD-ROM等の記録媒体やインターネット等の伝送媒体を介して配信できることはいうまでもない。

発明の効果

- [0018] 本発明のオーディオ信号符号化装置及び復号化装置によれば、原音信号を混入して得られる混入信号から原音信号を近似する再生信号を分離するための補助情報を生成する際に、人間の音源の方向知覚の特性、拡がり感、及び距離感に関する特徴をもちいることによって、聴感上、違和感がない程度に信号の分離を達成でき、かつ非常に小さな補助情報を生成することが可能となる。
- [0019] また、マルチチャンネルの原音信号から、前記混入信号として、左右2チャンネルのダウンミクス信号を得るように構成すれば、2チャンネル信号の再生系を有するスピーカやヘッドホンでの再生においては、補助情報を処理することなく、前記ダウンミクス信号を復号するだけで、低演算量で高音質なステレオ再生が可能となる。

図面の簡単な説明

- [0020] [図1]図1は、本発明の実施の形態に係るオーディオ信号符号化装置の機能的な構成の一例を示すブロック図である。
- [図2]図2は、チャンネル情報が示すリスナーと音源の位置関係の一例を表す図である。
- [図3]図3は、補助情報生成部の構成の一例を示す機能ブロック図である。
- [図4]図4(A)及び(B)は、知覚方向推定値の量子化に用いられるテーブルの一例を模式的に表す図である。
- [図5]図5(A)及び(B)は、信号間レベル差及び信号間位相差の量子化に用いられるテーブルの一例を模式的に表す図である。
- [図6]図6は、補助情報生成部の構成の他の一例を示す機能ブロック図である。
- [図7]図7は、信号間相関度、信号間類似度、及び拡がり感推定値の量子化に用いられるテーブルの一例を模式的に表す図である。
- [図8]図8は、補助情報生成部の構成のさらに他の一例を示す機能ブロック図である。
- [図9]図9は、本発明の実施の形態に係るオーディオ信号復号化装置の全体構成の

機能的な構成の一例を示すブロック図である。

[図10]図10は、信号分離処理部の構成の一例を示す機能ブロック図である。

符号の説明

[0021]	102	混入信号復号化部
	103	信号分離処理部
	105	第1出力信号
	106	第2出力信号
	201	第1入力信号
	202	第2入力信号
	203	混入信号符号化部
	204	補助情報生成部
	205	補助情報
	206	混入信号情報
	207	チャンネル情報
	303	信号間レベル差算出部
	304	信号間位相差算出部
	305	知覚方向推定部
	306	符号化部
	401	信号間相関度算出部
	402	拡がり感推定部
	403	符号化部
	502	距離感推定部
	503	符号化部
	702	補助情報
	704	混入信号復号化部
	705	復号化方法切り替え部
	706	信号間情報復号部
	707	信号合成部

発明の実施するための最良の形態

[0022] 以下、本発明の実施の形態について、図面を参照しながら説明する。

[0023] (オーディオ信号符号化装置)

図1は、本発明のオーディオ信号符号化装置の機能的な構成の一例を示すブロック図である。このオーディオ信号符号化装置は、外部から与えられる第1入力信号201及び第2入力信号202を符号化して、混入信号情報206を得ると共に、外部から与えられるチャンネル情報207によって示される各チャンネルの音声の再生位置の配置の関係に応じて異なる符号化方法を用いて補助情報205を得る装置であり、混入信号符号化部203及び補助情報生成部204からなる。

[0024] 混入信号情報206及び補助情報205は、後述するオーディオ信号復号化装置によって、第1入力信号201及び第2入力信号202それぞれを近似する信号に復号される情報であり、チャンネル情報207は、復号されるそれぞれの信号がリスナーから見てどの方向から再生されるかを表す情報である。

[0025] 図2は、信号再生用の音源とリスナーとの間の位置関係の一例を示す図である。この例は、5チャンネル再生を行う場合に各チャンネルの音源となるスピーカそれぞれのリスナーから見た配置方向を示している。例えば、フロントLチャンネルスピーカ及びフロントRチャンネルスピーカが、リスナーの正面からそれぞれ左右30° の方向に配置されることが示されている。この2つのスピーカは、ステレオ再生にも用いられる。

[0026] チャンネル情報207は、例えば、フロントLチャンネルスピーカ及びフロントRチャンネルスピーカから再生されるべき音声を符号化することを、具体的に、リスナーの正面を0° として反時計回りに+30° (フロントLチャンネルスピーカ)及び-30° (フロントRチャンネルスピーカ)といった音源の配置角度によって表す。また、実用的には30° といった細かな角度情報ではなく、各チャンネルの音源の配置角度を予め定めた上で、単にフロントLチャンネル、フロントRチャンネルといったチャンネル名によって同じことを表すこともできる。

[0027] チャンネル情報207は、どのチャンネルの音声を符号化するかを知っている外部の装置から適切に、本オーディオ信号符号化装置に与えられる。

[0028] 一つの典型例として、第1入力信号201及び第2入力信号202としてステレオの原

音信号が与えられ、そこからモノラルの混入信号と補助情報とを生成する場合には、フロントLチャンネル及びフロントRチャンネルを示すチャンネル情報207が与えられる。

- [0029] 別の典型例として、5チャンネルの原音信号から左右2チャンネルの混入信号を生成する際に、第1入力信号201及び第2入力信号202としてそれぞれフロントLチャンネル及びリアLチャンネルの信号を与えられ、そこから左チャンネルの混入信号と補助情報とを生成する場合には、フロントLチャンネル及びリアLチャンネルを示すチャンネル情報207が与えられる。
- [0030] 再び図1を参照して、第1入力信号201、および、第2入力信号202は、混入信号符号化部203と補助情報生成部204へと入力される。混入信号符号化部203では、ある一定の予め定められた方法によって第1入力信号201、および、第2入力信号202を加算することによって混入信号を生成し、その混入信号を符号化して得られた混入信号情報206を出力する。この符号化には、適宜周知の技術を用いることができるが、一例としては、背景技術の項で述べたAACなどの方法を用いてもよい。
- [0031] 補助情報生成部204では、第1入力信号201、第2入力信号202、混入信号符号化部203で作られた混入信号、及び混入信号情報206から、チャンネル情報207を用いて補助情報205を生成する。
- [0032] ここで、補助情報205は、混入信号から、混入前の原音信号である第1入力信号201及び第2入力信号202それぞれに聴感上できるだけ近い信号を分離するための情報である。補助情報205を用いて、混入信号から混入前の第1入力信号201と、第2入力信号202と完全に同じ信号を分離できてもよいし、聞いたところ差がないような程度の信号を分離できるのもよい。聞いて差があったとしても、その補助情報が信号分離のための情報である限り、本発明の範疇に含まれる。
- [0033] 補助情報生成部204は、チャンネル情報207を用いることによって、聴感上違和感のない程度の信号を分離可能な補助情報を、少ない情報量で生成する。そのために、補助情報生成部204は、チャンネル情報207に応じて、補助情報を符号化する方法、具体的には符号化の量子化精度を切り替える。
- [0034] 以下、補助情報生成部204のいくつかの実施例について、詳細に説明する。

[0035] (実施例1)

実施例1に係る補助情報生成部について図3から図5を用いて説明する。

[0036] 図3は、実施例1の補助情報生成部の機能的な構成を示すブロック図である。

[0037] 実施例1の補助情報生成部は、第1入力信号201及び第2入力信号202から、チャンネル情報207に応じて異なる符号化がなされた補助情報205Aを生成する部であり、信号間レベル差算出部303、信号間位相差算出部304、知覚方向推定部305、及び符号化部306からなる。

[0038] 補助情報205Aは、信号間レベル差算出部303で算出される信号間レベル差、信号間位相差算出部304で算出される信号間位相差、知覚方向推定部305で算出される知覚方向推定値のうちの少なくとも一つを量子化し符号化して得られる情報である。

[0039] 第1入力信号201と、第2入力信号202とは、信号間レベル差算出部303、および、信号間位相差算出部304へ入力される。

[0040] 信号間レベル差算出部303では、第1入力信号201と、第2入力信号202との信号のエネルギー差を算出する。エネルギー差を算出する場合において、信号を複数の周波数バンドに区切って、その各々について算出してもよいし、全帯域で1つ算出してもよい。また、算出する時間単位も特に限定しない。エネルギー差を表現する方法としても、オーディオ表現でよく用いられる指数関数値、たとえばdBとして差を表現してもよいし、必ずしも限定しない。

[0041] 信号間位相差算出部304では、第1入力信号201と第2入力信号202との信号間の相互相関を算出し、その相互相関値をもとに、その値が大きくなる位相差を算出する。このような位相差算出方法は、当該技術者には公知である。また、必ずしも相互相関値の最大値をあたえる位相を位相差にしなくてもいい。それは、相互相関はデジタル信号をもとに算出する場合、離散値であるがゆえに、位相差についても離散値で得られるからで、その解決手段としては、相互相関値の分布をもとに、補間により推定される位相差であってもよい。

[0042] 信号間レベル差算出部303から出力として得られる信号間レベル差、信号間位相差算出部304から出力として得られる信号間位相差、およびチャンネル情報207は、

知覚方向推定部305へ入力される。

- [0043] 知覚方向推定部305は、前記チャンネル情報207と、信号間レベル差算出部303から出力として得られる信号間レベル差、信号間位相差算出部304から出力として得られる信号間位相差から、リスナーが知覚する音像の方向を推定する。
- [0044] 一般に、2個のスピーカから音響信号を提示した際にリスナーが知覚する方向は、2個のスピーカの配置と、2ch信号のレベル差および位相差で決定されることが知られている(「空間音響」, イェンス ブラウエルト, 森本政之, 後藤敏幸 編著, 鹿島出版会(1986), あるいは, "Spatial Hearing: The Psychophysics of Human Sound Localization", revised edition, MIT Press, 1997)。知覚方向推定部305は、例えばこの知見に基づいて、リスナーによる音像の知覚方向を推定し、その推定結果を表す知覚方向推定値を符号化部306へ出力する。
- [0045] 符号化部306は、信号間レベル差、信号間位相差、及び知覚方向推定値のうちの少なくとも一つを、前述のチャンネル情報207と知覚方向推定値に応じて異なる精度で量子化し、さらに符号化を実施して得た補助情報205Aを出力する。
- [0046] 従来、リスナーの知覚弁別特性について、次のことが知られている。一般に、リスナーの知覚弁別特性は正面方向に対して左右対称であり、正面方向の知覚弁別特性が敏感で、フロントLチャンネル方向(もしくはフロントRチャンネル方向)に向かうにしたがって、鈍感になる傾向がある。また一般に、リスナーの知覚弁別特性は正面方向から反時計回りに背面方向まで前後非対称であり、正面方向の知覚弁別特性が敏感で、バックチャンネルに向かうにしたがって、鈍感になる傾向がある。
- [0047] 符号化部306は、このことを考慮して、知覚方向推定部305から得られた知覚方向推定値が、知覚弁別特性が敏感な方向を示す場合には信号間レベル差、信号間位相差、及び知覚方向推定値を細かく量子化し、逆に、知覚弁別特性が鈍感な方向を示す場合には敏感な方向を示す場合に比べて、荒く量子化する。
- [0048] つまり、符号化部306は、チャンネル情報207が、フロントLチャンネルとRチャンネルを示す場合は、知覚方向に関して左右方向に対称な量子化を実施し、フロントLチャンネルとバックLチャンネルを示す場合は、知覚方向に関して前後方向に非対称な量子化を実施する。

- [0049] このような量子化精度の切り替えを行うために、符号化部306は、一例として、入力値を量子化値へ変換する複数のテーブルを予め保持しておき、その中からチャンネル情報207に応じた一つを用いる。
- [0050] 図4は、符号化部306に予め保持される、知覚方向推定値の量子化に用いられるテーブルの一例を模式的に表す図である。何れのテーブルも知覚方向推定値の量子化点の一例を表しており、(A)はフロントLチャンネル及びフロントRチャンネル用の一例であり、(B)はリアLチャンネル及びフロントLチャンネル用の一例である。
- [0051] 符号化部306は、チャンネル情報207がフロントLチャンネル及びフロントRチャンネルを示す場合、図4(A)のテーブルに基づいて、知覚方向推定値を、知覚弁別特性が比較的敏感な正面方向の近傍では細かく量子化し、知覚弁別特性が比較的鈍感な左右方向に向かうにつれて荒く量子化する。
- [0052] また、チャンネル情報207がリアLチャンネル及びフロントLチャンネルを示す場合、図4(B)のテーブルに基づいて、知覚方向推定値を、知覚弁別特性が比較的敏感な正面方向の近傍では細かく量子化し、知覚弁別特性が比較的鈍感な背面方向に向かうにつれて荒く量子化する。
- [0053] 図5は、信号間レベル差及び信号間位相差の量子化に用いられるテーブルの一例を模式的に表す図である。何れのテーブルも所定の正規化がなされた信号間レベル差及び信号間位相差の量子化点の一例を表しており、(A)はフロントLチャンネル及びフロントRチャンネル用のテーブルの一例であり、(B)はリアLチャンネル及びフロントLチャンネル用のテーブルの一例である。
- [0054] 符号化部306は、チャンネル情報207がフロントLチャンネル及びフロントRチャンネルを示す場合、図5(A)のテーブルに基づいて、知覚方向推定値が知覚弁別特性の比較的敏感な正面方向の近傍を示す場合には、信号間レベル差及び信号間位相差を細かく量子化し、知覚方向推定値が知覚弁別特性の比較的鈍感な左右方向に向かうにつれて、信号間レベル差及び信号間位相差を荒く量子化する。
- [0055] また、チャンネル情報207がリアLチャンネル及びフロントLチャンネルを示す場合、図5(B)のテーブルに基づいて、知覚方向推定値が知覚弁別特性の比較的敏感な正面方向の近傍を示す場合には、信号間レベル差及び信号間位相差を細かく量子

化し、知覚方向推定値が知覚弁別特性の比較的鈍感な背面方向に向かうにつれて、信号間レベル差及び信号間位相差を荒く量子化する。

[0056] なお、図4及び図5のテーブルは何れも、本発明の特徴であるチャンネル情報207に応じて符号化方法を切り替えるための構成の一具体例を示したものであり、量子化点の分布を図示される内容に限定することを意図していない。チャンネル情報207が、リアLチャンネル及びリアRチャンネルを示す場合など、リスナーの知覚弁別特性を反映した量子化点の他の分布を表すテーブルを用いる場合も本発明に含まれる。

[0057] このテーブルを切り替える構成の他に、例えば量子化の関数や、符号化の手続きそのものを切り替えることによって、チャンネル情報207に応じた符号化方法の切り替えを実現することも考えられる。

[0058] 以上説明したように、符号化部306は、チャンネル情報207と知覚方向推定部305から得られる知覚方向推定値とにもとづいて、リスナーの音像の知覚方向に関する弁別能力が反映された量子化精度（つまり、正面方向には細かく、左右から背面方向に向かうにつれて荒くなる量子化精度）を決定し、信号間レベル差、信号間位相差、及び知覚方向推定値のうちの少なくとも一つを量子化し符号化する。

[0059] これにより、量子化精度を切り替えない場合に比べてより少ない情報量で表された補助情報を得ることが可能となる。

量子化精度の決定においては、音源が静止している場合の聴覚心理モデルに基づいて量子化テーブルや、量子化の関数を作成し量子化を実施してもよいし、実際の音源において、音像が移動することを考慮して、その音像の移動スピードや、量子化対象としている周波数帯域の特性に応じて、量子化精度を変えてもよい。特に時間分解能を適応的に変えることで、音源が静止している場合のモデルにあてはめて量子化し符号化することが可能となる。

[0060] このように構成された符号化方法を用いれば、人間の音の知覚方向の特性に基づいた符号化を実施でき、効率よく符号化を実施できる。

[0061] (実施例2)

実施例2に係る補助情報生成部について図6および図7を用いて説明する。

[0062] 図6は、実施例2の補助情報生成部の機能的な構成を示すブロック図である。

- [0063] 実施例2の補助情報生成部は、第1入力信号201及び第2入力信号202から、チャンネル情報207に応じて符号化された補助情報205Bを生成する部であり、信号間相関度算出部401、拡がり感推定部402、及び符号化部403からなる。
- [0064] ここで、補助情報205Bは、信号間相関度算出部401で算出される信号間相関度、信号間類似度、拡がり感推定部402で算出される拡がり感推定値のうちの少なくとも一つを量子化し符号化して得られる情報である。
- [0065] 第1入力信号201と、第2入力信号202とは、信号間相関度算出部401へ入力される。
- [0066] 信号間相関度算出部401では、第1入力信号201と、第2入力信号202との信号の相互相関値と各々の入力信号をもとに信号間の類似度(コヒーレンス)を、例えば次に示す式1
- (式1)
- $$ICC = \Sigma (x*(y+\tau)) / (\Sigma x*x \cdot \Sigma y*y)^{0.5}$$
- などに従って算出する。
- [0067] τ は、両耳間の位相ずれを補正するための項であり、当業者には公知である。
- [0068] 類似度を算出する場合において、信号を複数の周波数バンドに区切って、その各々について算出してもよいし、全帯域で1つ算出してもよい。また、算出する時間単位も特に限定しない。
- [0069] 信号間相関度算出部401から出力として得られる信号間の類似度、およびチャンネル情報207は、拡がり感推定部402へ入力される。
- [0070] 拡がり感推定部402は、前記チャンネル情報207と、信号間相関度算出部401から出力として得られる信号間の類似度から、リスナーが知覚する音像の拡がりの程度を推定する。ここでは、リスナーが知覚する音像の拡がりの程度は、心理量としての大きさや小ささを適宜数値化して表現されるものとする。
- [0071] 一般に、音の拡がり感は、リスナーの両耳に入力される音響信号の音圧レベルと両耳間相関度で説明できることが知られている(特許第3195491号および特許第3214255号)。ここで、両耳間相関度(DICC)とチャンネル間相関度(ICC)は、次の式2

で表される関係にある。

[0072] (式2) $DICC = ICC \times Clr$

ここで、ClrはHlとHrの相関度であり、Hlはスピーカなどの音源からリスナーの左耳までの伝達関数、Hrはスピーカなどの音源からリスナーの右耳までの伝達関数である。ここで、リスニングルームなどのようにスピーカ配置が左右対称である場合Clrは1とみなせるので、音像の拡がり感は信号間相関度と音圧レベルで予測することができる。拡がり感推定部402は、例えばこの知見に基づいて、リスナーが感じる音の拡がり感を推定し、その推定結果を表す拡がり感推定値を符号化部403へ出力する。

[0073] 符号化部403は、信号間相関度、信号間類似度、及び拡がり感推定値のうちの少なくとも一つを、前述のチャンネル情報207に応じて異なる精度で量子化し、さらに符号化して得た補助情報205Bを出力する。

[0074] 従来、両耳間相関度が同じでも、リスナーに到来する直接音の方向がリスナーの正面ではない場合は、直接音が正面から到来する場合と比較して拡がり感が減少することが知られている(“Relation between Auditory Source Width in Various Sound Fields and Degree of Interaural Cross-Correlation”, M. Morimoto, K. Iida, and Y. Furue, Applied Acoustics, 38 (1993) 291-301)。

[0075] これは、音声の再生が、フロントLチャンネルとフロントRチャンネルから行われる場合に比べて、フロントLチャンネルとバックLチャンネルから行われる場合には、再生音の拡がり感に対するリスナーの弁別能力が劣ることを意味している。

[0076] 符号化部403は、このことを考慮して、チャンネル情報207が、フロントLチャンネル及びフロントRチャンネルを示す場合と、フロントLチャンネル及びバックLチャンネルを示す場合とで、異なる精度で量子化を実施する。

[0077] このような量子化精度の切り替えを行うために、符号化部403は、一例として、入力値を量子化値へ変換する複数のテーブルを予め保持しておき、その中からチャンネル情報207に応じた一つを用いる。

[0078] 図7は、符号化部403に予め保持される、信号間相関度、信号間類似度、及び拡がり感推定値の量子化に用いられるテーブルの一例を模式的に表す図である。何れのテーブルも、所定の正規化がなされた信号間相関度、類似度、拡がり感推定値の

量子化点の一例を表しており、(A)はフロントLチャンネル及びフロントRチャンネル用の一例であり、(B)はリアLチャンネル及びフロントLチャンネル用の一例である。

[0079] 符号化部403は、チャンネル情報207がフロントLチャンネル及びフロントRチャンネルを示す場合、図7(A)のテーブルに基づいて、信号間相関度、信号間類似度、及び拡がり感推定値を、比較的細かく量子化し、チャンネル情報207がリアLチャンネル及びフロントLチャンネルを示す場合、図7(B)のテーブルに基づいて、信号間相関度、信号間類似度、及び拡がり感推定値を、比較的荒く量子化する。

[0080] 以上説明したように、符号化部403は、チャンネル情報207に基づいて、リスナーの拡がり感に対する弁別能力が反映された量子化精度(つまり、正面方向には細かく、左右から背面方向へ向かうにつれて荒くなる量子化精度)を決定し、決定された量子化精度で信号間相関度、信号間類似度、及び拡がり感推定値のうちの少なくとも一つを量子化し符号化する。

[0081] このように構成された符号化方法を用いれば、人間の音像の拡がり感の特性に基づいた符号化を実施でき、効率よく符号化を実施できる。

[0082] (実施例3)

実施例3に係る補助情報生成部について図8を用いて説明する。

[0083] 図8は、実施例3に係る補助情報生成部の機能的な構成を示すブロック図である。

[0084] 実施例3の補助情報生成部は、第1入力信号201、第2入力信号202から、チャンネル情報207に応じて符号化された補助情報205Cを生成する部であり、信号間相関度算出部401、距離感推定部502、及び符号化部503からなる。

[0085] ここで、補助情報205Cは、信号間相関度算出部401で算出される信号間相関度、信号間類似度、距離感推定部502で算出される距離感推定値のうちの少なくとも一つを量子化し符号化して得られる情報である。

[0086] 第1入力信号201と、第2入力信号202とは、信号間相関度算出部401へ入力される。

[0087] 信号間相関度算出部401では、第1入力信号201と、第2入力信号202との信号の相互相関値と各々の入力信号をもとに信号間の類似度(コヒーレンス)を前述した式1などに従って算出する。

- [0088] 類似度を算出する場合において、信号を複数の周波数バンドに区切って、その各々について算出してもよいし、全帯域で1つ算出してもよい。また、算出する時間単位も特に限定しない。
- [0089] 信号間相関度算出部401から出力として得られる信号間の類似度、およびチャンネル情報207は、距離感推定部502へ入力される。
- [0090] 距離感推定部502は、前記チャンネル情報207と、信号間相関度算出部401から出力として得られる信号間の類似度から、リスナーが知覚する音像の距離感の程度を推定する。ここでは、リスナーが知覚する音像の距離感の程度は、心理量として感じられる遠さや近さを適宜数値化して表現されるものとする。
- [0091] 従来、リスナーが知覚する音像の距離感と、前述した式1によって算出される信号間相関度算出部401からの出力値(類似度)の正負の符号との間に関係があることが知られている。このことは、「2チャンネル音響信号の相関係数と音像の質:日本音響学会誌 39巻4号(1983)」(黒住幸一ら)に記載されている。距離感推定部502は、例えばこの知見に基づいて、リスナーが知覚する音像の距離感を推定し、その推定結果を表す距離感推定値を符号化部503へ出力する。
- [0092] 符号化部503は、信号間相関度、信号間類似度、及び距離感推定値のうちの少なくとも一つを、前述のチャンネル情報207に応じて異なる精度で量子化し、さらに符号化して得た補助情報205Cを出力する。
- [0093] 再生音の距離感に関しても、音声の再生が、フロントLチャンネルとフロントRチャンネルから行われる場合と、フロントLチャンネルとバックLチャンネルから行われる場合とで、リスナーの弁別能力が異なることが予想される。
- [0094] 符号化部503は、このことを考慮して、チャンネル情報207が、フロントLチャンネル及びフロントRチャンネルを示す場合と、フロントLチャンネル及びバックLチャンネルを示す場合とで、異なる量子化を実施する。
- [0095] このような量子化精度の切り替えを行うために、符号化部503は、一例として、入力値を量子化値へ変換する複数のテーブルを予め保持しておき、その中からチャンネル情報207に応じた一つを用いる。このようなテーブルには、図7で説明したテーブルと同様のテーブルが使えるため、ここでは詳細な説明を省略する。

- [0096] 以上説明したように、符号化部503は、チャンネル情報207に基づいて、リスナーが感じる音像までの距離感に関する弁別能力が反映された量子化精度（つまり、正面方向には細かく、左右から背面方向に向かうにつれて荒くなる量子化精度）を決定し、決定された量子化精度で、信号間相関度、信号間類似度、及び距離感推定値のうちの少なくとも一つを量子化し符号化する。
- [0097] このように構成された符号化方法を用いれば、人間の音像の距離感の特性に基づいた符号化を実施でき、効率よく符号化を実施できる。
- [0098] （実施例4）
- 実施例4のオーディオ信号符号化装置は、実施例1と実施例2、および実施例3のオーディオ信号符号化装置を組み合わせるものである。
- [0099] 実施例4のオーディオ信号符号化装置は、図3、図6、および図8に示される構成の全てを兼ね備え、2つの入力信号から信号間レベル差、信号間位相差、信号間相関度（類似度）を算出し、チャンネル情報に基づいて、知覚方向および拡がり感、および距離感を推定し、チャンネル情報に応じて、量子化方法や量子化テーブルを切り替えて、符号化を実施する。
- [0100] なお、実施例4では、実施例1から実施例3のうち、いずれかの2つを組み合わせてもよい。
- [0101] （オーディオ復号化装置）
- 図9は、本発明のオーディオ信号復号化装置の機能的な構成の一例を示すブロック図である。このオーディオ信号復号化装置は、前述のオーディオ信号符号化装置によって生成される混入信号情報206、補助情報205、及びチャンネル情報207から、原音信号を近似する第1出力信号105及び第2出力信号106を復号する装置であり、混入信号復号化部102及び信号分離処理部103からなる。
- [0102] 本発明は、オーディオ信号符号化装置からオーディオ信号復号化装置へ、混入信号情報206、補助情報205、及びチャンネル情報207を移送する具体的な方法を限定しないが、一例としては、混入信号情報206、補助情報205、及びチャンネル情報207が、一つの放送ストリームに多重化されて送信され、オーディオ信号復号化装置は、その放送ストリームを受信して多重分離することによって、混入信号情報206、補

助情報205、及びチャンネル情報207を得るとしてもよい。

- [0103] また、例えば、混入信号情報206、補助情報205、及びチャンネル情報207が、記録媒体に蓄積されており、オーディオ信号復号化装置は、その記録媒体から、混入信号情報206、補助情報205、及びチャンネル情報207を読み出すとしてもよい。
- [0104] なお、チャンネル情報207については、オーディオ信号符号化装置とオーディオ信号復号化装置の間でその規定値や規定の順番を予め取り決めておくことによって、伝送を省略することも考えられる。
- [0105] 混入信号復号化部102は、符号化データの形式で表される混入信号情報206を、オーディオ信号の形式へと復号し、復号されたオーディオ信号を信号分離処理部103へ出力する。混入信号復号化部102は、前述のオーディオ信号符号化装置における混入信号符号化部203の逆変換を行うものであり、例えば、混入信号符号化部203が混入信号情報206をAACに従って生成する場合には、混入信号復号化部102もまたAACに定められる逆変換を行って前記オーディオ信号を得る。前記オーディオ信号の形式は、時間軸上の信号形式、周波数軸上の信号形式、及び時間と周波数の両軸で表現される形式などから選択されるものとし、本発明ではその形式を限定しない。
- [0106] 信号分離処理部103は、混入信号復号化部102から出力されるオーディオ信号から、補助情報205及びチャンネル情報207に基づいて、第1出力信号105と第2出力信号106とを生成し、出力する。
- [0107] 以下、信号分離処理部103の詳細について説明する。
- [0108] 図10は、この実施例に係る信号分離処理部103の機能的な構成を示すブロック図である。
- [0109] 信号分離処理部103は、チャンネル情報207に応じて異なる復号化方法を用いて補助情報205を復号化し、その復号結果を用いて第1出力信号105及び第2出力信号106を生成する部であり、復号化方法切り替え部705、信号間情報復号部706、及び信号合成部707からなる。
- [0110] 復号化方法切り替え部705は、チャンネル情報207を入力されると、そのチャンネル情報207に基づいて、信号間情報復号部706に対して復号化方法の切り替えを

指示する。

- [0111] 信号間情報復号部706は、復号化方法切り替え部705からの指示に従って切り替えた復号化方法を用いて、補助情報702を信号間情報に復号化する。この信号間情報は、実施例1から実施例3で説明した、信号間レベル差、信号間位相差、信号間相関度などである。信号間情報復号部706は、オーディオ信号符号化装置における符号化部と同様に、量子化点を表すテーブルを切り替えることによって復号化方法を切り替えることができる。他にも、例えば量子化の逆関数や、復号化の手続きそのものを切り替えることによって復号化方法を切り替えてもよい。
- [0112] 信号合成部707は、混入信号復号化部704の出力信号であるオーディオ信号から、前記信号間情報によって表される信号間レベル差、信号間位相差、及び信号間相関度を持った第1出力信号105と第2出力信号106とを生成する。この生成には、前記オーディオ信号を複製して得た2つの信号に、前記信号間レベル差の半分をそれぞれ逆方向に与え、前記信号間位相差の半分をそれぞれ逆方向に与え、さらに、レベル差及び位相差を付与後の2つの信号を前記信号間相関度に応じて混入する、といった周知の方法を適宜用いればよい。
- [0113] このように構成された復号化方法を用いれば、チャンネル情報を反映した効率のよい復号化方法が達成でき、高音質な複数信号を得ることが可能となる。
- [0114] また、この復号化方法は、1チャンネルのオーディオ信号を2チャンネルのオーディオ信号にすることだけでなく、nチャンネルのオーディオ信号を、nより多いチャンネルのオーディオ信号に生成することに用いることが可能である。たとえば、2チャンネルのオーディオ信号から6チャンネルのオーディオ信号を得る場合や、1チャンネルのオーディオ信号から6チャンネルのオーディオ信号を得る場合などでも有効である。
- 産業上の利用可能性
- [0115] また、本発明のオーディオ信号復号化装置、オーディオ信号符号化装置、及びそれらの方法は、オーディオ符号化されたビットストリームを伝送するシステム、例えば、放送コンテンツの伝送システム、DVDやSDカードなどの蓄積媒体に音声情報を記録、再生するシステム、携帯電話に代表される通信機器にAVコンテンツを伝送するシステムなどに利用できる。また、インターネット上でやりとりされる電子データとして、

オーディオ信号を伝送するシステムにおいても利用できる。

請求の範囲

- [1] 複数のチャンネルそれぞれの原音信号を、前記原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とに符号化するオーディオ信号符号化装置であって、
前記原音信号を混入して得た混入信号を符号化することによって前記混入信号情報を生成する混入信号符号化手段と、
前記原音信号に基づいて前記特徴量を算出すると共に、各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる符号化方法を決し、前記決定された符号化方法を用いて前記算出された特徴量を符号化することによって前記補助情報を生成する補助情報生成手段と
を備えることを特徴とするオーディオ信号符号化装置。
- [2] 前記補助情報生成手段は、異なる量子化精度が得られる量子化点を定義する複数のテーブルを予め記憶しており、前記特徴量を、前記複数のテーブルのうちの前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つによって定義される量子化点に量子化することによって符号化することを特徴とする請求項1に記載のオーディオ信号符号化装置。
- [3] 前記補助情報生成手段は、前記原音信号間のレベル差及び位相差の少なくとも一方を算出して前記特徴量とすることを特徴とする請求項1に記載のオーディオ信号符号化装置。
- [4] 前記補助情報生成手段は、前記原音信号間のレベル差及び位相差の両方を算出し、前記算出されたレベル差及び位相差に基づいて、リスナーが知覚すると推定される音像の方向を算出して前記特徴量とすることを特徴とする請求項3に記載のオーディオ信号符号化装置。
- [5] 前記補助情報生成手段は、リスナーの正面方向から左右対称に設けられる量子化点を定義する第1のテーブルと、リスナーの左方向から前後非対称に設けられる量子化点を定義する第2のテーブルとを予め記憶しており、前記チャンネル情報がリスナーの左前方と右前方とを示す場合には、前記特徴量を前記第1のテーブルによって

定義される量子化点に量子化することによって符号化し、前記チャンネル情報がリスナーの左前方と左後方とを示す場合には、前記特徴量を前記第2のテーブルによって定義される量子化点に量子化することによって符号化する

ことを特徴とする請求項3に記載のオーディオ信号符号化装置。

- [6] 前記補助情報生成手段は、前記原音信号間の類似度を算出して前記特徴量とする

ことを特徴とする請求項1に記載のオーディオ信号符号化装置。

- [7] 前記補助情報生成手段は、前記原音信号間の相互相関値、又はその絶対値を、前記類似度として算出する

ことを特徴とする請求項6に記載のオーディオ信号符号化装置。

- [8] 前記補助情報生成手段は、前記算出された類似度に基づいて、リスナーが知覚すると推定される音像の拡がり感及び距離感の少なくとも一方を算出して前記特徴量とする

ことを特徴とする請求項6に記載のオーディオ信号符号化装置。

- [9] 複数のチャンネルそれぞれの原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とを、前記複数のチャンネルそれぞれの再生信号に復号化するオーディオ信号復号化装置であって、

各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられ、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる復号化方法を決定する復号化方式切り替え手段と、

前記決定された復号化方法を用いて前記補助情報を前記特徴量に復号化する信号間情報復号手段と、

前記混入信号情報と前記復号化された特徴量とを用いて、前記複数のチャンネルそれぞれの再生信号を生成する信号合成手段と

を備えることを特徴とするオーディオ信号復号化装置。

- [10] 前記補助情報は、前記特徴量を、異なる量子化精度が得られる量子化点を定義する複数のテーブルのうち前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つによって定義される量子化点に量子化することによって符号化さ

れており、

前記信号間情報復号手段は、前記複数のテーブルを予め記憶しており、前記複数のテーブルのうちの前記チャンネル情報によって示される各再生位置の配置の関係に応じた一つを用いて、前記補助情報を前記特徴量に復号化する

ことを特徴とする請求項9に記載のオーディオ信号復号化装置。

- [11] 前記特徴量は、前記原音信号間のレベル差、位相差、及び、リスナーが知覚すると推定される音像の方向の少なくとも一つを表し、

前記信号間情報復号手段は、リスナーの正面方向から左右対称に設けられる量子化点を定義する第1のテーブルと、リスナーの左方向から前後非対称に設けられる量子化点を定義する第2のテーブルとを予め記憶しており、前記チャンネル情報がリスナーの左前方と右前方とを示す場合には、前記第1のテーブルを用いて前記補助情報を前記特徴量に復号化し、前記チャンネル情報がリスナーの左前方と左後方とを示す場合には、前記第2のテーブルを用いて前記補助情報を前記特徴量に復号化する

ことを特徴とする請求項10に記載のオーディオ信号復号化装置。

- [12] 前記特徴量は、前記原音信号間のレベル差、位相差、類似性、及び、リスナーが知覚すると推定される音像の方向、拡がり感、並びに距離感の少なくとも一つを表すことを特徴とする請求項9に記載のオーディオ信号復号化装置。

- [13] 前記信号合成手段は、前記特徴量が、前記原音信号間のレベル差、位相差、及び類似性の少なくとも一つを表す場合に、前記混入信号情報によって表される音声信号に前記特徴量に応じたレベル差、位相差、及び類似性を与えることによって、前記再生信号を生成する

ことを特徴とする請求項12に記載のオーディオ信号復号化装置。

- [14] 複数のチャンネルそれぞれの原音信号を、前記原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とに符号化するオーディオ信号符号化方法であって、

前記原音信号を混入して得た混入信号を符号化することによって前記混入信号情報を生成する混入信号符号化ステップと、

前記原音信号に基づいて前記特徴量を算出すると共に、各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる符号化方法を決定し、前記決定された符号化方法を用いて前記算出された特徴量を符号化することによって前記補助情報を生成する補助情報生成ステップと

を含むことを特徴とするオーディオ信号符号化方法。

- [15] 複数のチャンネルそれぞれの原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とを、前記複数のチャンネルそれぞれの再生信号に復号化するオーディオ信号復号化方法であって、
各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる復号化方法を決定する復号化方式切り替えステップと、

前記決定された復号化方法を用いて前記補助情報を前記特徴量に復号化する信号間情報復号ステップと、

前記混入信号情報と前記復号化された特徴量とを用いて、前記複数のチャンネルそれぞれの再生信号を生成する信号合成ステップと

を含むことを特徴とするオーディオ信号復号化方法。

- [16] 複数のチャンネルそれぞれの原音信号を、前記原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とに符号化するための、コンピュータ実行可能なプログラムであって、

前記原音信号を混入して得た混入信号を符号化することによって前記混入信号情報を生成する混入信号符号化ステップと、

前記原音信号に基づいて前記特徴量を算出すると共に、各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられると、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる符号化方法を決定し、前記決定された符号化方法を用いて前記算出された特徴量を符号化することによって前記補助情報を生成する補助情報生成ステップと

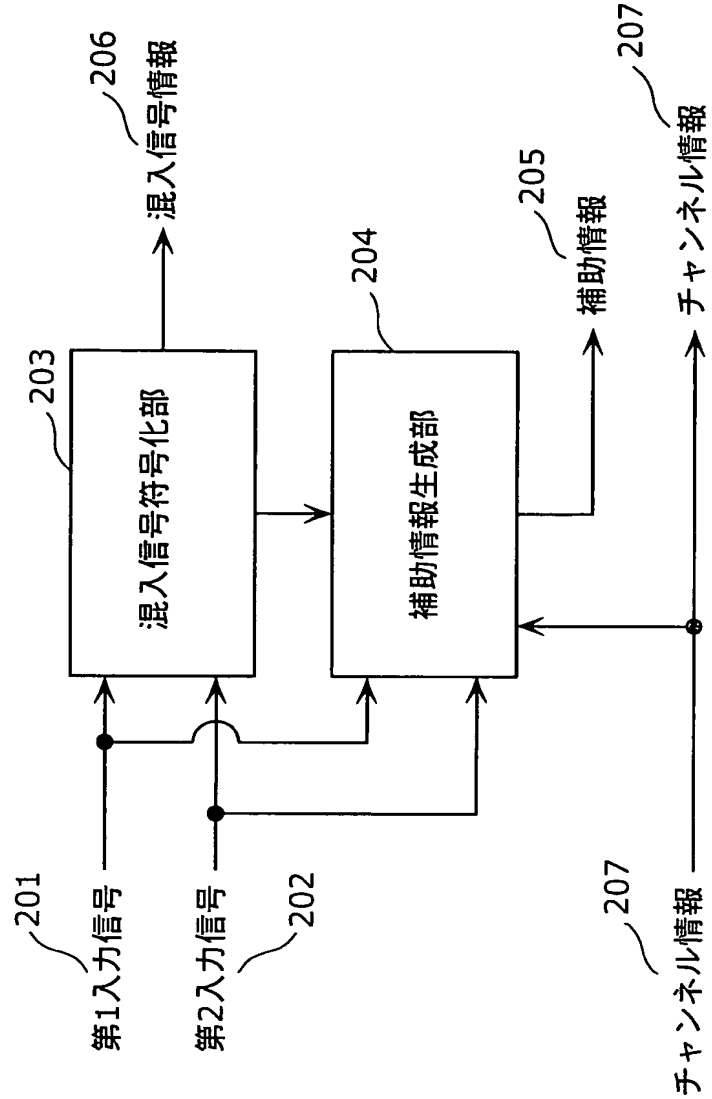
をコンピュータに実行させることを特徴とするプログラム。

- [17] 複数のチャンネルそれぞれの原音信号の全体的な特徴を表す混入信号情報と、個々の原音信号間の関係に基づく特徴量を表す補助情報とを、前記複数のチャンネルそれぞれの再生信号に復号化するためのコンピュータ実行可能なプログラムであって、
- 各チャンネルの音声のリスナーからみた再生位置を示すチャンネル情報が与えられ、与えられたチャンネル情報によって示される各再生位置の配置の関係に応じて異なる復号化方法を決定する復号化方式切り替えステップと、
- 前記決定された復号化方法を用いて前記補助情報を前記特徴量に復号化する信号間情報復号ステップと、
- 前記混入信号情報と前記復号化された特徴量とを用いて、前記複数のチャンネルそれぞれの再生信号を生成する信号合成ステップと
- をコンピュータに実行させることを特徴とするプログラム。
- [18] 請求項16及び請求項17の少なくとも一方に記載のプログラムを格納していることを特徴とするコンピュータ読み取り可能な記録媒体。

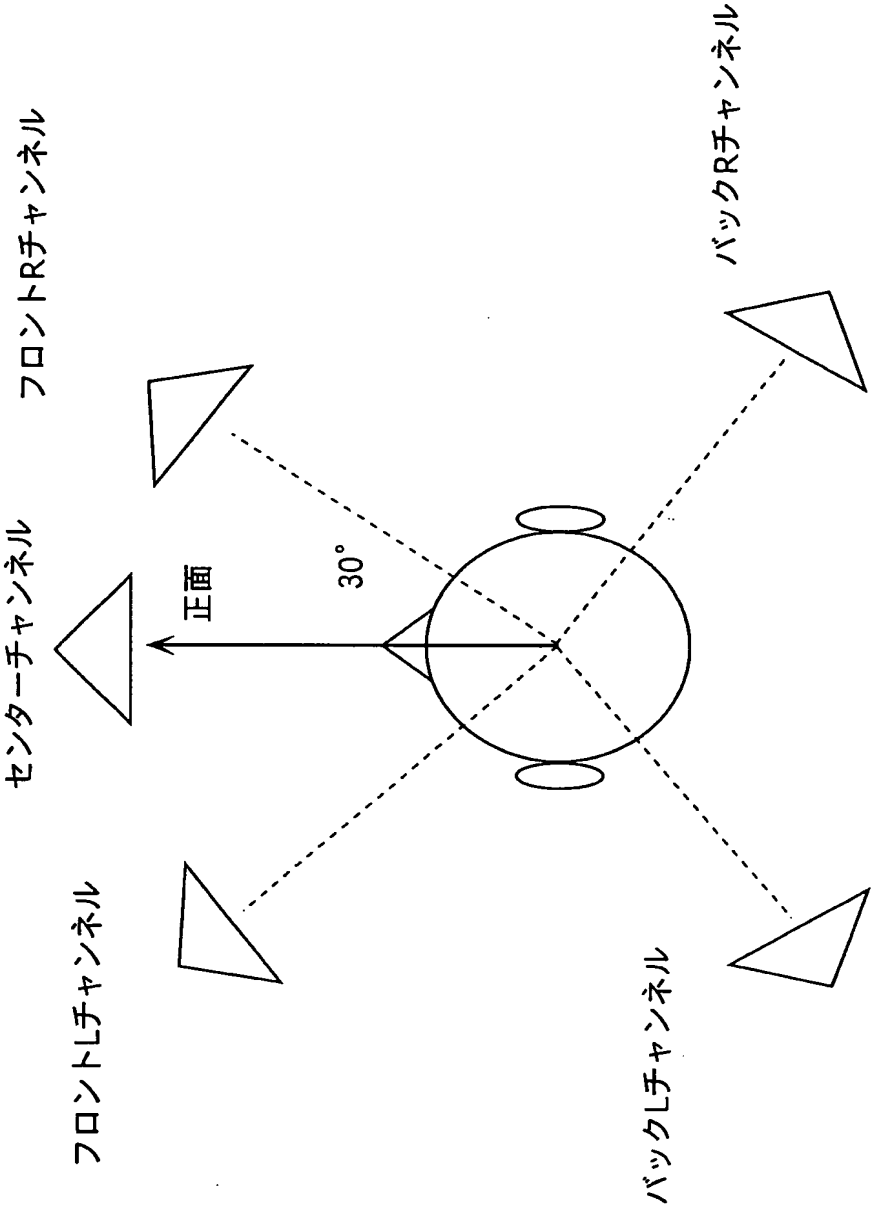
要 約 書

混入信号符号化部203および補助情報生成部204からなるオーディオ信号符号化装置であって、ある一定の予め定められた方法によって加算した混入信号を生成し、その混入信号を符号化し、混入信号情報206を出力する混入信号符号化部203と、第1入力信号201と、第2入力信号202、および、混入信号符号化部203で作られた混入信号と混入信号情報206を用いて補助情報205を生成する補助情報生成部204とを備える。補助情報生成部204は、人間の音源の方向知覚の特性や、拡がり感、および距離感に関する特徴を用いて、補助情報205を効率よく量子化する。

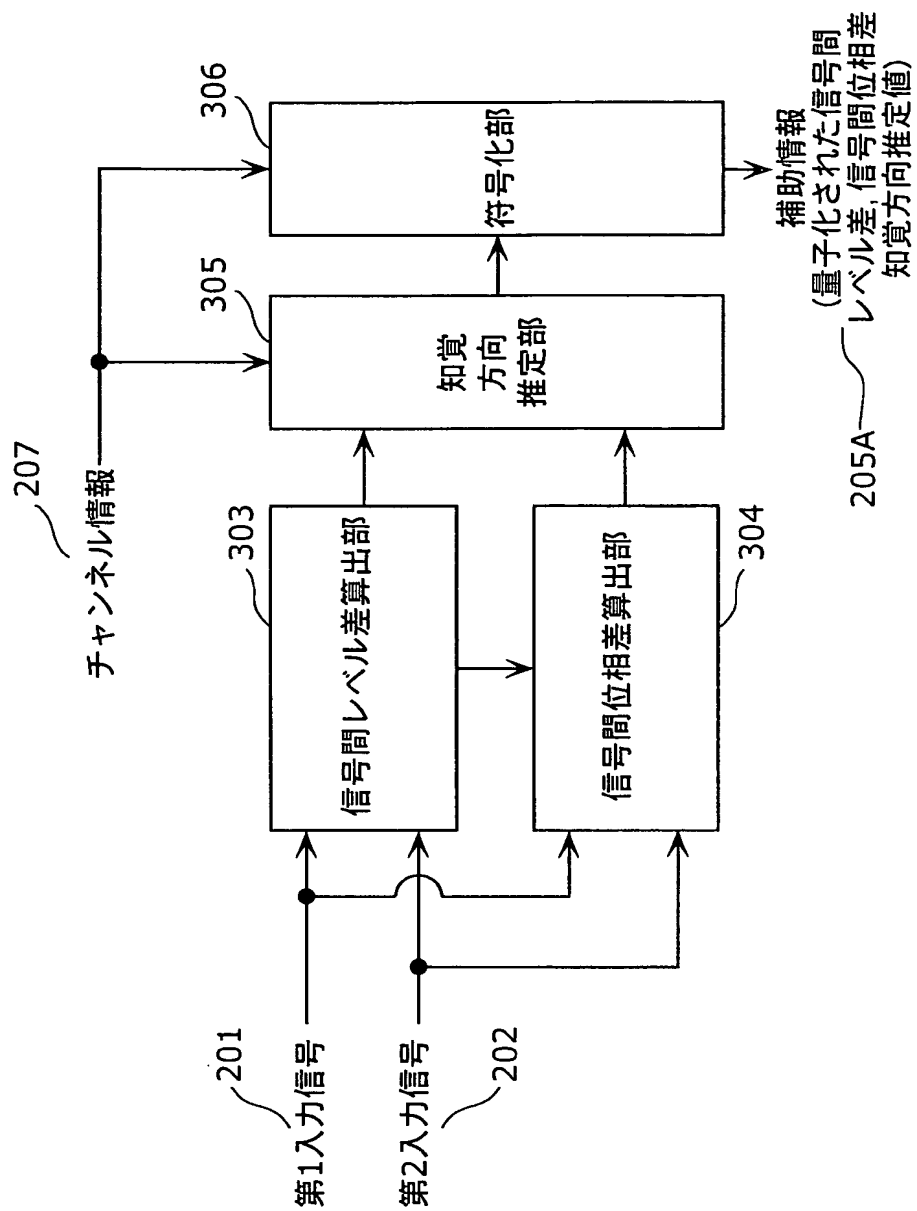
[図1]



[図2]



[図3]



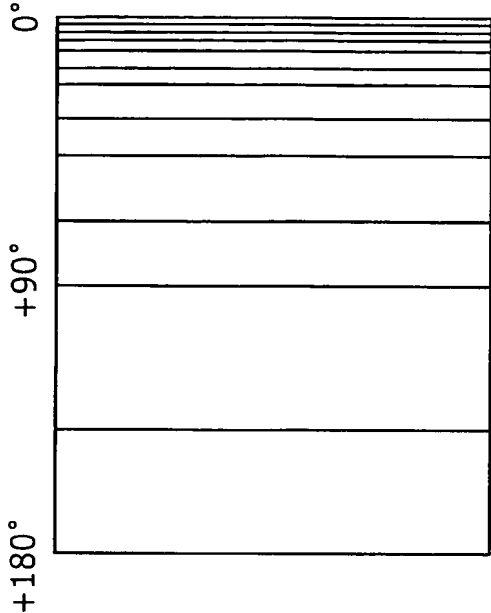
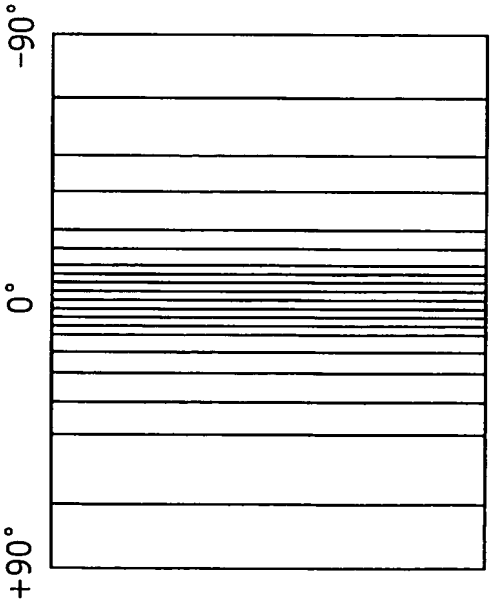
[図4]

(A) フロントLチャンネル-フロントRチャンネル用

(B) リアルチャンネル-フロントLチャンネル用

知覚方向推定値の量子化点

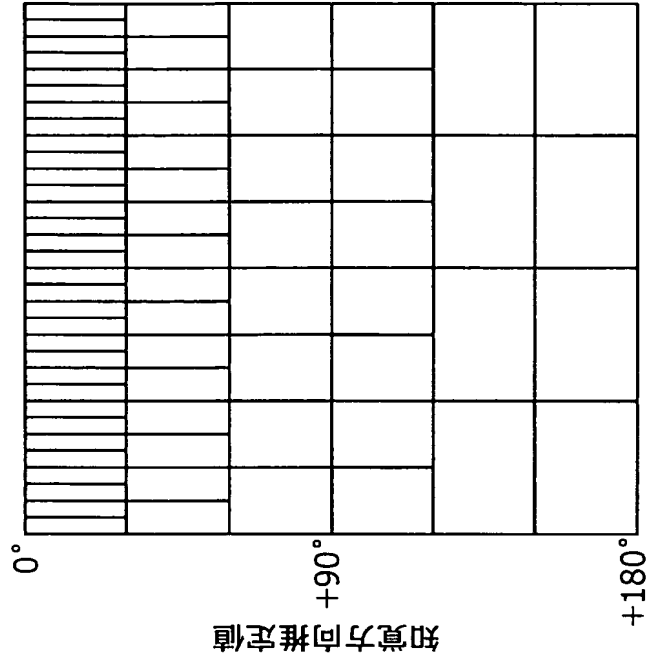
知覚方向推定値の量子化点



[図5]

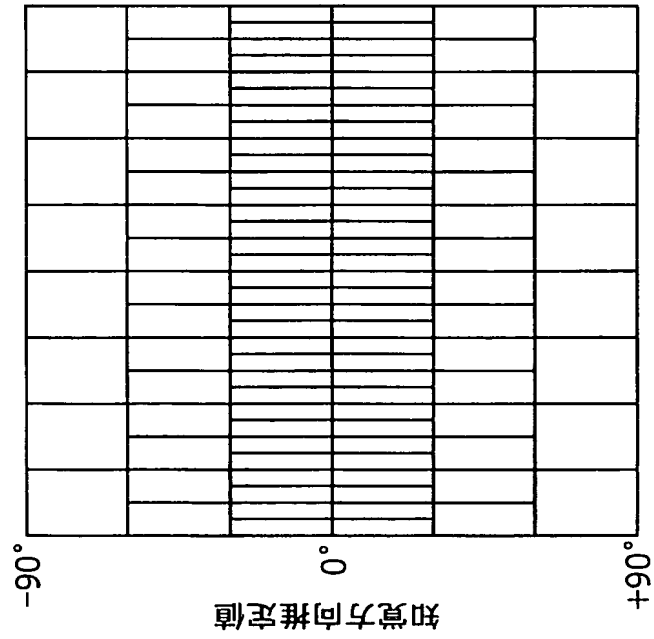
(B) リアルチャンネル-フロントLチャンネル用

信号間レベル差, 信号間位相差の量子化点

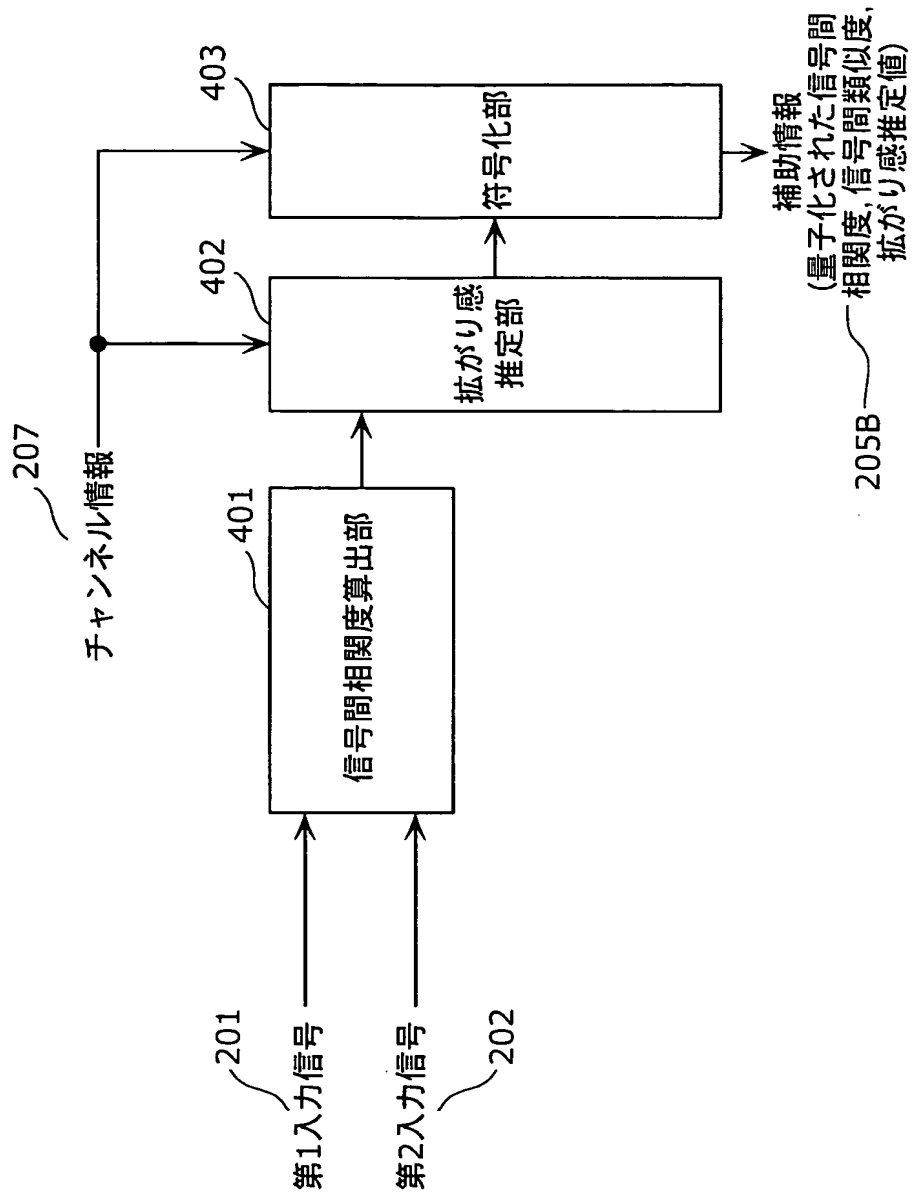


(A) フロントLチャンネル-フロントRチャンネル用

信号間レベル差, 信号間位相差の量子化点

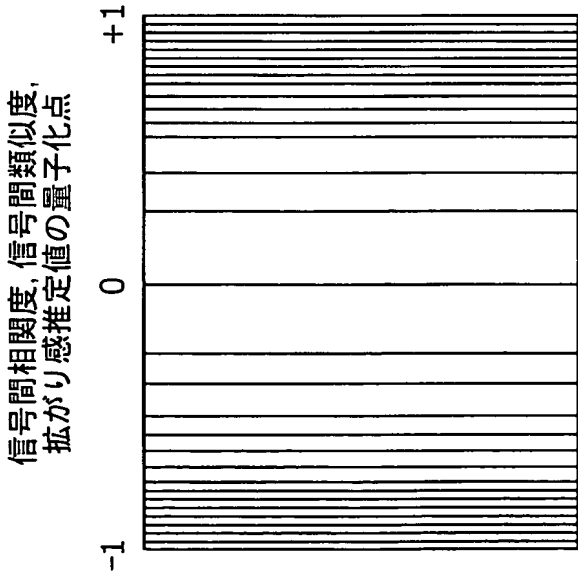


[図6]

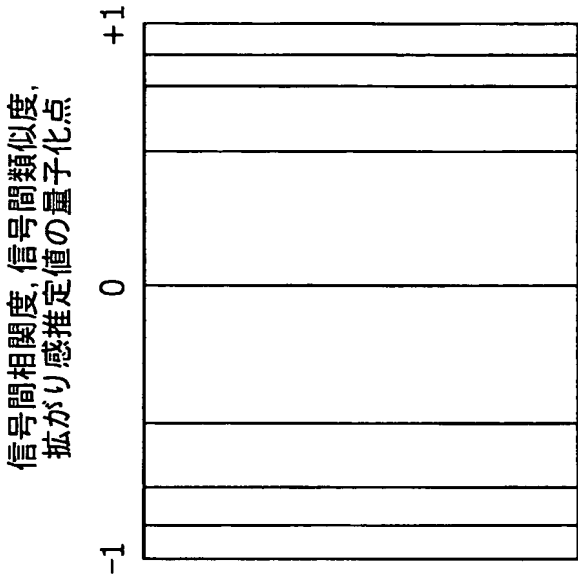


[図7]

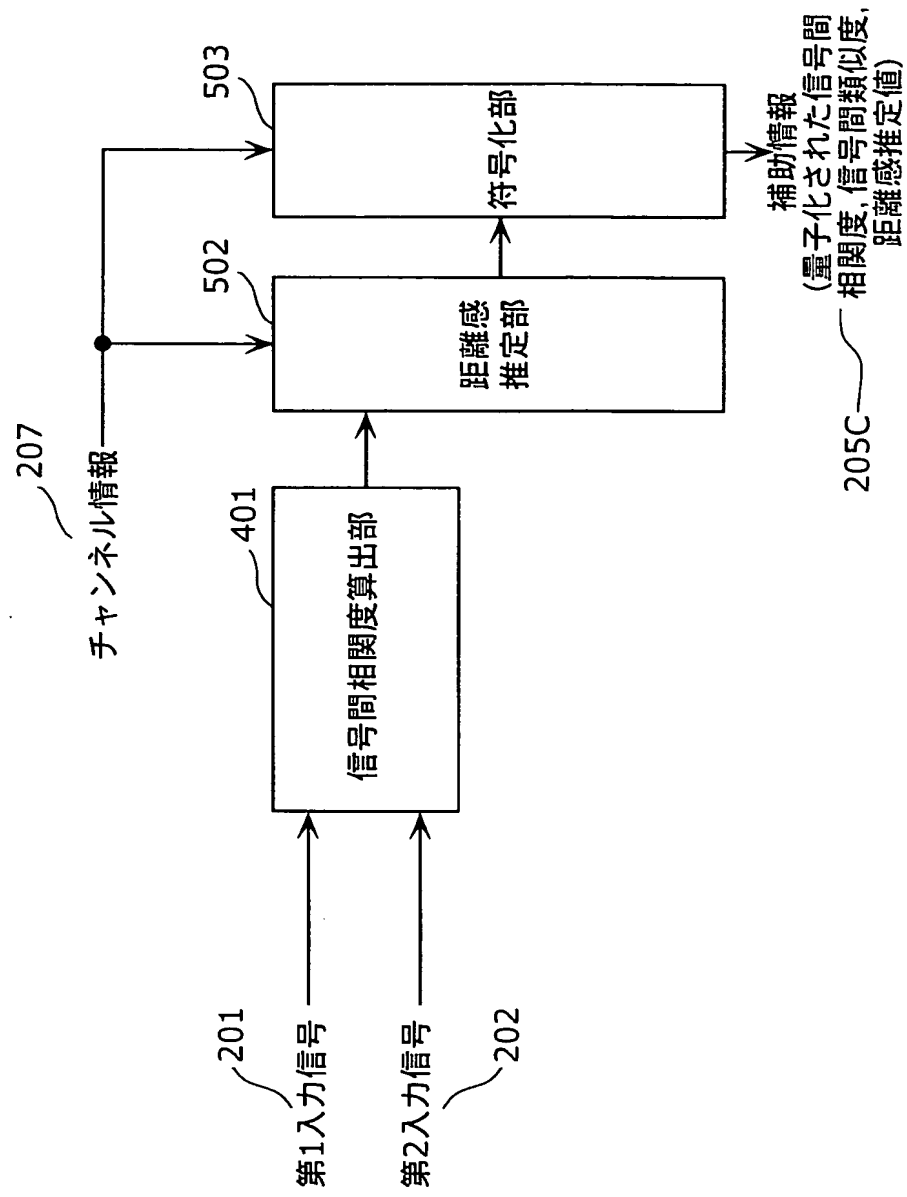
(A) フロントLチャンネル-フロントRチャンネル用



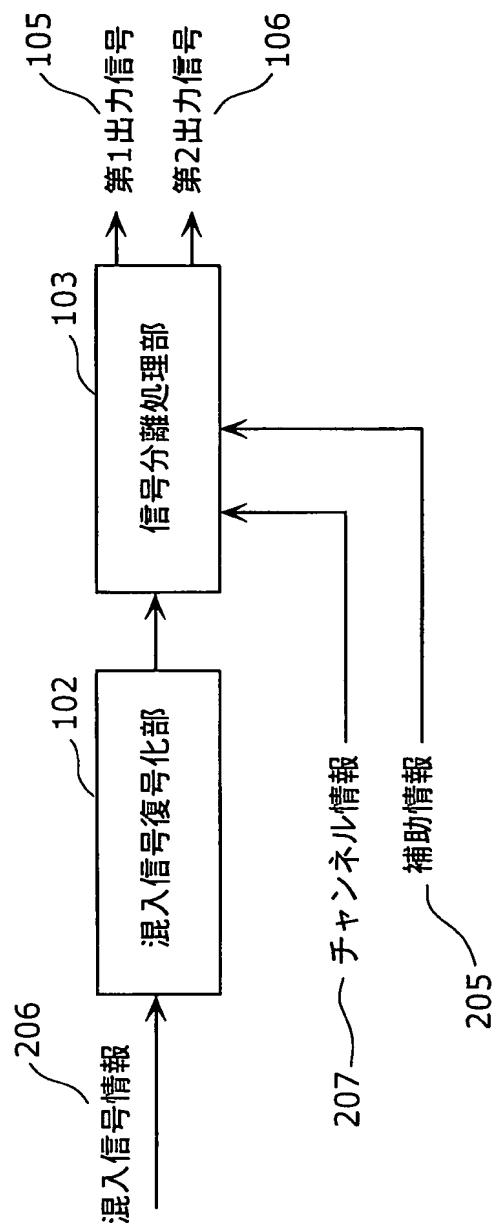
(B) リアルチャンネル-フロントLチャンネル用



[図8]



[図9]



[図10]

